

**UNITED STATES AIR FORCE
ARMSTRONG LABORATORY**

**REDISCOVERING LEARNING:
Acquiring Expertise in Real World
Problem Solving Tasks**

Sherrie P. Gott

**HUMAN RESOURCES DIRECTORATE
Cognition and Performance Division
7909 Lindbergh Drive
Brooks Air Force Base, TX 78235-5352**

19980526 154

April 1998

DTIC QUALITY INSPECTED &

Approved for public release; distribution is unlimited.

NOTICES

Publication of this paper does not constitute approval or disapproval of the ideas or findings. It is published in the interest of scientific and technical information (STINFO) exchange.

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely Government-related procurement, the United States Government incurs no responsibility or any obligation whatsoever. The fact that the Government may have formulated or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication, or otherwise in any manner construed, as licensing the holder or any other person or corporation; or as conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

The Office of Public Affairs has reviewed this paper, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This paper has been reviewed and is approved for publication.

SHERRIE P. GOTT
Project Scientist

R. BRUCE GOULD
Technical Director

Please notify AFRL/HEOP (ARS), 2509 Kennedy Circle, Brooks AFB, TX 78235-5118, if your address changes, or if you no longer want to receive our technical reports. You may write or call the STINFO Office at DSN 240-3877 or commercial (210) 536-3877.

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.				
1. AGENCY USE ONLY-(Leave blank)		2. REPORT DATE May 1998		3. REPORT TYPE AND DATES COVERED Sep 1993 - Dec 1995
4. TITLE AND SUBTITLE Rediscovering Learning: Acquiring Expertise in Real World Problem Solving Tasks			5. FUNDING NUMBERS PE - 63227F PR - 2949 TA - 00 WU - 03	
6. AUTHOR(S) Sherrie P. Gott				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Armstrong Laboratory (AFMC) Human Resources Directorate 7909 Lindbergh Drive Brooks Air Force Base TX 78235-5352			8. PERFORMING ORGANIZATION REPORT NUMBER AL/HR-TP-1997-0009	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)			10. SPONSORING/MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES Invited Keynote Address presented at (re)Forming Post-Compulsory Education Training Conference, Brisbane, Queensland, Australia, Dec 94, published in <i>Australian and New Zealand Journal of Vocational Educational Research</i> , Vol 3, number 1.				
12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) The importance of continuous learning in high-tech work settings is being rediscovered as industry and the military services react to external forces such as increasingly complex and rapidly changing equipment systems as well as highly competitive product/service markets. Competitiveness in turn dictates a leaner, downsized workforce for the private sector, and diminished defense spending has resulted in dramatic losses of personnel in the Armed Forces. Those who remain are expected to do more, and yet, performance demands routinely override training opportunities. Moreover on-the-job training that follows either the traditional master-apprentice behavioral model or relies heavily on didactic instruction is typically impractical or ineffective. An alternative learning-oriented approach that accelerates skill acquisition in high-tech jobs is described here. With this approach cognitive performance models provide both the input to instruction and the desired criterion performance to be attained. The instructional medium is an intelligent tutoring system. A knowledge elicitation approach called the PARI cognitive task analysis methodology is described, along with the cognitive models of performance yielded by this analysis. The performance models in turn inform a coached apprenticeship practice environment embodied in an intelligent computer tutor. The system was recently evaluated in a controlled experiment at three geographically separated Air Force workcenters. Results reveal that the experimental group significantly accelerated their acquisition of problem solving skills when compared to a matched control group; moreover, their newly acquired troubleshooting skills generalized to a novel equipment system. The dramatic learning effects are attributed to cognitive models as input to instruction, situated learning in a constructivist instructional environment, and the sociology surrounding the tutoring system whereby apprentices gradually become members of their community of practice.				
14. SUBJECT TERMS cognitive task analysis, cognitive performance models, practical learning experiences, cognitive skills, learning by doing, learning through reflection, intelligent computer tutors, skill generality			15. NUMBER OF PAGES 38	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UL	

CONTENTS

	Page
Introduction	1
Academic vs. Practical Learning Experiences	2
Cognitive Performance Models	3
Individual Differences and Other Influences	4
Downsizing Inevitably Means Expanded Job Responsibilities	5
Research Aimed at Rediscovering Learning for Real Word Problem Solving	6
Intelligent Computer Tutors Informed by Cognitive Models	10
Theory-Based Instruction	12
Method	13
Results	14
Discussion	17
Experience Equivalence Analyses	17
Cognitive Models as Input to Instruction	18
Situated Learning in a Constructivist Environment	19
Coaching	20
Reflective Follow-up (RFU) Activities	22
Sociology of the Learning Environment	23
Quality of Knowledge Representations -- Runnable Mental Models	24
Presence of a Top-Level Goal Structure	25
Multilevel Structures to Transfer.	26

SYNTHESIS AND SUMMATION	26
REFERENCES	29

LIST OF FIGURES

Figure	Page
1 Overview and Illustration of PARI Methodology	7
2 Cognitive Skills Architecture	8
3 Top-Level Mental Model of Avionics Equipment System	11
4 Actual Avionics Test Station	11
5 Pre to Posttest Changes: Sherlock 2 Tests	15
6 Sherlock 2 Posttest Scores Across Groups	15
7 Frankenstation Test Scores Across Group.	16
8 Example of Weak Performer's Mental Model.	21
9 Example of Strong Performer's Mental Model	21

LIST OF TABLES

Table	Page
1 Posttest Measures of Troubleshooting Proficiency (Sherlock 2)	14
2 Posttest Measures of Transfer (Frankenstation)	16
3 Effect Size for Posttest Measures	17

PREFACE

I am indebted to the hundreds of F15 avionics technicians at Langley AFB VA, Eglin AFB FL, Luke AFB AZ, and Nellis AFB NV, whose expertise and conscientious dedication to their mission as aircraft maintainers have made this work possible. More specifically, I am especially indebted to the three F15 avionics technicians who have worked in our laboratory as active members of the research team: MSgt Mark Gallaway, MSgt Ron Kane, and TSgt Dennis Collins. Thanks is due as well to the leadership and staff of HQ TAC/LG (ACC/LG) for their long-term, stalwart support of the Basic Job Skills (BJS) R&D Program.

The quality of the science in this and every other BJS product has been immeasurably enhanced by contributions from our troika of eminent scientific advisors: Dr. Robert Glaser (chair), University of Pittsburgh; Dr. David Kieras, University of Michigan; and Dr. Robert Linn, University of Colorado.

Finally, we dedicate this report and others in the BJS series to Gen. Henry Viccellio, Jr., without whose vision and unwavering support this work would not have reached fruition and made the difference it has in the Air Force maintenance community.

The research reported in this paper was conducted in full partnership with the Learning Research and Development Center, University of Pittsburgh, Prof. Alan Lesgold, Principal Investigator. The opinions expressed herein are those of the author and do not necessarily reflect those of the Air Force. Correspondence concerning this paper should be addressed to Dr. Sherrie Gott, Armstrong Laboratory, AL/HRMD, 7909 Lindbergh Drive, Brooks AFB TX, 78235-5352.

REDISCOVERING LEARNING: Acquiring Expertise in Real World Problem Solving Tasks

INTRODUCTION

Why do we find ourselves in the position of hearkening a call to **rediscover learning**? Why has the importance of learning the skills central to a modern workforce been allowed to wane?

One hypothesis suggests a widening gap between what schools produce as educated individuals and what the real world needs as competent workers. A society of unprecedented technological sophistication is evolving, a society that now demands considerable specialized knowledge and competencies from its citizenry. But where are such learning experiences provided? In the past, these are the capabilities that have been considered the purview of out-of-school training programs. Today, however, the boundaries are less clear. A striking discontinuity thus exists between what the schools supply and what the real world demands as skilled performers. The importance of **learning** on the job is being **rediscovered** as we attempt to meet the demands.

Historically, many occupations have included **practical apprenticeships** -- or internships as part of their training experience. Teachers, physicians, attorneys, psychotherapists, bankers, as well as craftspeople in technical fields such as carpentry, mechanics, and plumbing have traditionally been required to complete practice-centered internships. In theory, these programs should provide a fruitful source of learning experiences that are practice- (or performance-) oriented. In reality, one finds only limited instances where such practice-oriented training directly targets the high-tech, knowledge-rich tasks that have changed the face of the modern workplace.

One explanation for this phenomenon (in the USA) has been the rise of the vocational education movement, which has resulted in internships shifting back into the purview of school-based learning. This shift unfortunately has too often meant that hands-on practice and coaching were replaced by traditional, didactic instruction (Resnick, 1987). "As the ideology of expanded schooling took hold and as the nature of the workplace changed, we gave up opportunities for learning in the workplace in favor of school-based vocational education" (Resnick, 1987, p. 17). In turn, conventional procedures for educational measurement, which focused on declarative knowledge acquisition, were applied. To take the point a bit further, it is not unreasonable to suggest that the absence of progressive, skill-oriented **assessment procedures** may in fact be related to the demise of practice-centered training. Norman Frederiksen (1984) offered an

insightful account of such precedents in military training, where changing the test meant instructional reform. After all, students learn that on which they will be tested, and teaching facts such as Ohm's Law and Kirchoff's Voltage and Current Laws is certainly easier to test on a standard examination than the student's capability to explain how the laws affect electrical functionality in a faulted circuit. It may not be hyperbole, then, to claim that truly effective experiential training may never become a reality until valid assessment is in place to shape curriculum.

Given the poor fit between models of school-based learning and society's need for skilled workers, it is not surprising that the industrial and military sectors have significantly increased their own training initiatives. American businesses are spending billions of dollars to establish and operate what have been termed "in-house corporate colleges" (Mitchell, 1987). It is estimated that as much as \$40 billion a year is being spent to deliver instruction to eight million workers. For example, Motorola -- a maker of electronic products -- recently spent two years retraining its mechanical engineers for jobs as electronic engineers, because those skills are more highly demanded by information-age technologies (Mitchell, 1987).

Technology-driven instructional needs are similarly affecting military technical training. Increases and modifications in technical instruction are most often found in on-the-job training (OJT) environments, not in formal academic settings. In the U.S. military, the operational commands no longer count exclusively on the traditional models of school-based learning used by training commands to develop the technically skilled workers that are needed. Yet, even OJT programs generated by the operational community -- which clearly amount to a form of educational protest -- bear startling similarities to the very academic models of school house learning they reject. Often the only notable difference is that the factual knowledge that fills the curriculum is more specialized in character. Still missing, by and large, are practice-oriented learning experiences and, as one might expect, accompanying practical learning assessments. As a way of speculating why this is so, let me contrast performance-oriented learning with academic instructional models.

Academic vs. Practical Learning Experiences

School-based learning is directed toward the acquisition of theoretical knowledge and the development of general competency skills (Resnick, 1987). Glaser and others (Glaser, Lesgold, Lajoie, Eastman, Greenberg, Logan, Magone, Weiner, Wolf, & Yengo, 1985) have argued that to be effective, academic learning must target components of the developing proficiencies as informed by a cognitive theory of performance. Achievement components of interest might include:

- the structure and interrelatedness of declarative knowledge;
- the conditionalized nature of that knowledge, i.e., the degree to which it is associated with indicators of how and when it is to be appropriately used;
- mental models of task demands and task characteristics;

- théories held by students to explain certain phenomena; and
- automaticity of task subprocesses.

In contrast, practice-oriented instructional programs are directed not at general academic proficiencies, such as knowledge of a particular subject matter (e.g., physics) or general computational skills, but rather toward some specific criterion performance, such as computer programming or electronics troubleshooting. The level of achievement desired is beyond the initial stages of learning, where factual knowledge bases are constructed. Rather, emphasis is on the later stages of skill acquisition, where knowledge is proceduralized and procedures are in turn smoothed out via practice. Assessments of practical learning should therefore be capable of pinpointing weaknesses in the proceduralization process, i.e., deficiencies in the learner's successive approximations of mature practice as demonstrated by use of knowledge under appropriate conditions, guided by top-level goals, and by efficient task execution regulated by metacognitive skills.

Because the targeted level of achievement in practice-centered training is a specific criterion performance, the nature of that performance has considerable influence on the design of instruction and associated assessment. The more overt and observable the elements of the criterion performance, the more it lends itself to traditional forms of apprenticeship training and assessment. For example, a carpenter's apprentice can learn a great deal by following a conventional apprentice regimen of observing the master, executing a task with support and critique from the master, and then practicing extensively, soon becoming autonomous. Similarly, assessment can be accomplished by evaluating observable behaviors and products of the behaviors, using standards provided by expert craftspeople. However, practical learning experiences for modern workplace tasks are not so easily handled.

Criterion performances in the modern workplace are more mental and less physical; external behavioral elements are replaced by internalized cognitive structures, processing, and events. Learning through observation is significantly hampered, and as a consequence traditional instruction and assessment are generally impotent. No longer is it effective to focus on overt behaviors and observable end products. Rather, practical learning experiences and assessments for modern work environments must be targeted at the internalized processes and concepts that lie behind the successive approximations of expertise. An accepted principle of practice-oriented training is that testing and training should **mirror** the criterion performance. The same principle applies for modern cognitive apprenticeships. As with academic learning and assessment, a guiding theory of performance is clearly needed to inform the practical learning and assessment process.

Cognitive Performance Models

In conceiving of the sciences of the artificial, Simon (1981) offers a general theory of performance that is relevant here. Simon characterized human performance (and learning) as

moving across environments of varying complexity in pursuit of particular goals. With this guiding conception of performance, intelligent behavior can mean, among other things, steadily in complexity in recent decades as information system hardware and software have proliferated. Interacting with complex machines in one's job is now the rule for most workers, not the exception; however, the nature of intelligent performance in those interactions is still not well understood. As a result, characterization and subsequent measurement of worker performance are often misguided because of the vagueness surrounding what it means to be skilled in a complex technical domain.

Simon's model for mastering (and simplifying) one's environment notwithstanding, formal institutional attempts at simplifying the new high technology work environments consist mainly of thicker instructional manuals and technical documents. In other words, institutions have generated more "stuff" to accompany more technology. However, when we carefully analyze skilled performers to learn how they actually represent their environment internally and do their work, what we find are **not** detailed memorial replicas of dense technical data, but rather streamlined mental representations, or models, of the workings of the systems about which all the words are written. As skilled workers learn what their duties demand of them, they economically and selectively construct and refine their domain knowledge and procedural skill. That is one way they simplify their jobs. This suggests that the determinants of competence are not always revealed by the surface characteristics of either the worker's performance or the environment in which that performance takes place.

Complex job environments require deeper cognitive analyses that can ferret out the conceptions and reasoning processes that lurk behind observable behaviors. Recent work of cognitive psychologists interested in expert problem solving has advanced our understanding of reasoning and envisioning processes. For example, Larkin, McDermott, Simon, & Simon (1980) have shown that rich sets of schemata indexed by large numbers of patterns underlie the quickness of mind and insightful views of expert "intuitive" problem solvers, those who can fill in a sketchy representation with just the right pieces of information. Further, expert-novice studies have consistently shown that experts first reason **qualitatively** about a problem to understand causal relations before moving to quantitative formalisms (Chi, Feltovich, & Glaser, 1981; deKleer, 1984; Gitomer, 1984). In support of this thesis, students of the physical sciences have consistently shown deficiencies and misconceptions in their understanding of the underlying causal principles of the domain (di Sessa, 1982; Larkin, et al., 1980; McCloskey, Caramazza, & Green, 1980). And yet, the quantitative formalisms (such as Kirchoff's Voltage and Current Laws) still dominate scientific instruction.

Individual Differences and Other Influences

A complex work environment further complicates practical, performance-related training because of the inherent variability in human information processing. In a complex environment where so much is to be apprehended, encoded, and represented in memory, individual differences

in cognition assume considerable importance. The interactions between a worker's cognitive apparatus (including all-important prior knowledge) and the many features of the complex systems encountered in the workplace are considerable. Even when a group of people enters the workplace after an apparently uniform initial job training experience, each person brings his or her own set of conceptions about the domain just studied. In such situations, a cognitive analysis that examines complex human performance in depth may uncover uniformities as well as common misconceptions or bugs that affect the on-the-job learning processes. This can lead to better design and development of the kind of adaptive training that can significantly facilitate learning in complex domains.

There are also sources of influence that are external to the worker. First of all, there are the cultural pressures of the workplace. Particularly in the military, the apprentice is pressed to learn the job quickly in order to become a contributing member of the workforce as soon as possible. Typically, it is the cadre of apprentices who are relied on as the critical mass or core capability of a military operational unit. Simultaneously, however, the demand characteristics of a military unit can be quite severe, which is to say that putting planes in the air in the Air Force, for example, takes precedence over on-the-job training sessions. Learning the job quickly is thus frequently impeded because of the demand to get the work out at all costs. Opportunities to practice and refine skills are characteristically nonexistent, particularly for the worker of average skill or below.

Downsizing Inevitably Means Expanded Job Responsibilities

Another potent influence on learning and job competence comes as a consequence of one of the management philosophies currently favored by both the military and the private sector -- **downsizing**. In theory, by assigning broader responsibilities to a given class of worker, technicians somehow become transformed from specialists to generalists. In practice, such a policy necessitates dramatic changes in instructional practices in order for broader domain knowledge and more flexible reasoning skills to be acquired. Effective training to foster skill flexibility rarely accompanies downsizing and its baggage of broadened job responsibilities for remaining personnel.

Already, training practices have been weakened under the weight of complex subject matter and formidable workplace machines. An argument that has sometimes prevailed is that smarter machines mean reduced cognitive loads on workers and that consequently **less** training is required. Of course, machines capable of automating certain workplace tasks, i.e., the relatively easy portions of the jobs, do not in reality appreciably reduce the cognitive workload. Rather, what the machines do is take responsibility for the lower order or algorithmic tasks, reducing the apprentice to a passive observer who is called into action only when nondeterministic problem solving is required. In other words, the apprentice loses opportunities to **learn by doing** the routine workplace tasks but is expected to somehow acquire the ability to solve non-routine problems either when the machine breaks down or when the problem is beyond the machine's

capabilities. Advances in technology have unquestionably contributed to lost apprenticeship learning opportunities, while at the same time improving shop productivity on routine tasks.

The complex machines also pose logistics problems to the training community who have the formidable task of evaluating the increasingly complex workplaces to determine instructional goals. Tough questions about the fidelity of the training place vis-à-vis the workplace have been vigorously debated. The training place typically has low priority for the expensive machines that populate the workplace, an unfortunate consequence of which is that during training, hands-on learning opportunities are often replaced by theoretical abstractions that cannot be tied to concrete experience.

All of this translates into the following kind of scenario for the typical apprentice in today's workplace: initial technical training is customarily patterned on an academic model of teaching complex subject matter. Students are **told** about a work domain instead of receiving practice in it. The academically-trained apprentice is met at the workplace by high expectations and by demand characteristics that simultaneously increase the pressure to learn, while eliminating many of the needed learning opportunities. The implications for the interplay of cognitive apprenticeships, performance assessments, and **learning** in this context are enormous.

Research Aimed at Rediscovering Learning for Real World Problem Solving

The Air Force Human Resources Directorate of the Armstrong Laboratory recently completed a long-term program of research and development investigating issues of technical competence in modern Air Force workcenters. The Basic Job Skills (BJS) research program was directed at examining human problem solvers in real-world, machine-laden work environments -- our goal being to improve our understanding of the acquisition and flexibility of technical expertise -- in effect **to rediscover learning** and its role in the modern workplace. The premise is that with greater understanding, we can build better training and make personnel resource-related decisions that are more responsive to high-tech performance demands. Better training has the potential to accelerate the rate of complex skill acquisition, and further, if training focuses on the components of skill that are common to a range of complex tasks, it becomes possible to equip technicians with generalizable skills -- i.e., a technical flexibility that greatly increases the value of human capital in an era of rapid technological change and diminishing defense resources.

Our research strategy for addressing these issues unfolded in four stages of investigation, as follows: (a) during a pilot testing phase, we observed and interviewed F15 avionics maintenance personnel troubleshoot problems at varying levels of complexity and subsequently selected a separate sample of airmen to participate in a proof of concept experiment; (b) based on preliminary findings, we developed, tested, refined and applied a formal cognitive task analysis (CTA) methodology to elicit from experts the components of skill for this domain; (c) we

codified the output of the task analysis and then used it as input to the development of an intelligent tutoring system for avionics troubleshooting; and (d) we evaluated the tutor in a controlled experiment at operational flying wings.

The cognitive task analysis approach called the PARI method (Precursor - Action - Result - Interpretation) (Gott, 1987; Gott, Bennett, & Gillet, 1986; Hall, Gott, & Pokorny [in press]; Means & Gott, 1988) blends the processes of a thinking aloud protocol approach (Ericsson & Simon, 1993) with the inherent structure of diagnostic problem solving activity, i.e., represent the problem, formulate diagnostic plan, hypothesize, test, interpret results, hypothesize and test again. An overview of the PARI interview process and an example of a PARI node from an actual interview are shown in Figure 1. This methodology constitutes the lever by which we came to explicate expertise and rediscover the importance of practical learning experiences in complex domains, such as electronic troubleshooting.

The PARI structured interview technique is a cognitive task analysis approach designed to investigate diagnostic reasoning processes, but is generalizable to other complex cognitive tasks as well. It has been designed to produce detailed solution traces and accompanying elaborations (generated by solution rehashes) so that fine-grained protocol data are systematically collected under actual conditions of problem solving.

The PARI process unfolds in dyadic problem-solving sessions. Members of the dyad are typically two experts, one who poses a problem scenario and responds with Results, and one who attempts to solve the problem. A task analyst probes the problem solver for the reasons behind the Actions he elects to take and for his Interpretations of the Results of the Actions. The analyst also asks the problem solver to draw a representation of the equipment system, as he envisions it in the mind's eye, to illustrate each PARI node. In this way, the reasoning processes -- particularly top level goals and strategic planning -- are made explicit and thus become part of the cognitive model of performance that is being generated for instructional use. An example is provided below.

A distinguishing characteristic of the PARI approach is the emphasis on revealing knowledge and skill in the context of their use. This approach contrasts with decontextualized task analysis or knowledge engineering interviews where knowledge is abstracted and detached from the problem solving conditions under which it is normally applied.

Precursor: I want to see if the LRU resistor is good.

Action: Remove the cable from J12 of the LRU and ohm out the path through the LRU from pin 68 to pin 128.

Result: 1.56 Mohms

Interpretation: The problem isn't in the LRU. It's in the test station or test package.

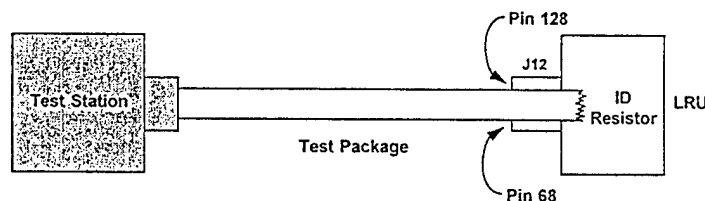


Figure 1. Overview and Illustration of PARI Methodology

In the third stage of the study, we used the PARI analytic tool to capture realistic problem solving performances across a range of troubleshooting tasks and human proficiency levels. The resultant data supplied us the means to induce a model of technical performance that has guided our development of the intelligent tutoring systems called Sherlock (1 and 2). (See detailed tutor description in next section.) An abstracted representation of the model is shown as a cognitive skills architecture in Figure 2. The power of the model can be easily illustrated: whether operating a word processor or diagnosing a faulty engine, the human performer in today's workplace is required to **select** and execute **procedures** to interact with an **object** to achieve a set of **goals**. The knowledge and processes that constitute that performance are (a) procedural (or how-to-do-it) knowledge; (b) declarative (domain) knowledge of the object (often called system or device knowledge); and (c) strategic (or how-to-decide-what-to-do-and-when) knowledge (Gott, 1989). With this decomposition, it is assumed that **procedural** and **device** knowledge are organized and deployed by mechanisms such as the goals, plans, and decision rules that comprise **strategic** knowledge. (Gott & Pokorny, 1987). This deployment capability serves a control function to enable what can be called **dynamic, opportunistic reasoning**. Ideally, this results in optimal solutions crafted in response to particular situations by applying just the right piece of knowledge at just the right time. This configuration as a model of technical performance posits that a top-level plan or strategy deploys pieces of system knowledge and procedural subroutines as needed and as driven by strategic decision factors such as time, effort, payoff, and resource efficiency. Troubleshooting is thereby represented as multilevel, complex decision making, which involves choices among various top-level and intermediate-level strategies, as well as among localized tactics. For the Sherlock tutors, this cognitive skills architecture was instantiated by the CTA process to yield codified models of avionics troubleshooting expertise at the level of detail necessary to inform instruction.

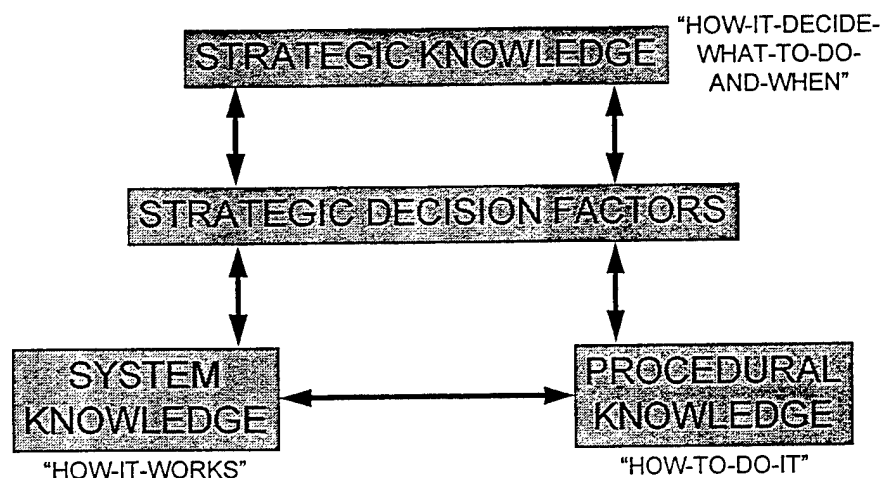


Figure 2. Cognitive Skills Architecture

Instances where instruction has failed to achieve close correspondence to complex (cognitive) performance models have been documented in a variety of academic and technical domains (Anderson, Boyle, Farrell, & Reiser, 1984; Greeno, 1978; Kieras, 1987). The typical outcome is suboptimal learning. When instruction falls short, it is often strategic control knowledge (vs system or procedural knowledge) that is most often overlooked. In cases where instructional materials have been evaluated against cognitive models, the instructional gaps that have been identified correspond to goal structure and other types of strategic knowledge, that is, the **glue** that gives coherence to the performance (Greeno, 1978; Anderson, 1984; Kieras, 1987). Strategic knowledge is arguably the least visible component of troubleshooting, and yet, support for the critical role of strategic knowledge in complex performances has been amassed in academic domains such as geometry (Greeno, 1978) as well as in practical domains such as text editing (Card, Moran, & Newell, 1983), computer programming (Anderson, Boyle, Farrell, & Reiser, 1984), simple device operation (Kieras, 1987), and electronic troubleshooting (Gott & Pokorny, 1987). Moreover, there is a large literature on the role of metacognitive knowledge in accomplishing planning, self-regulation, and performance control functions (see, for example, Brown, 1978; Brown, Bransford, Ferrara, & Campione, 1983).

Given the targeted models of expertise revealed by the cognitive task analysis, i.e., electronic troubleshooting as multi-component, complex decision making, we knew that to be effective instructionally, the learning environment had to be robust. (In the real world, expertise of this type takes 8 to 10 years to develop.) We adopted the following multifaceted principle of learning as the foundation for instructional design: in complex diagnostic tasks, mental models (system knowledge) as well as procedural and strategic knowledge are **constructed** as students interact with the full context of the work environment, practicing shop procedures and fault isolation operations in response to realistic troubleshooting scenarios. Trainees receive support from coaching, which they access as needed. To culminate the process, learners **reflect upon** their solutions considering their strengths, diagnosing their weaknesses, and contemplating **model** solutions of Masters. In short, this principle calls for a situated, supported learning environment, which we have termed **coached apprenticeship**. The intelligent tutoring system, though a relatively new instructional innovation at the time of Sherlock 1 (1984), nonetheless seemed sufficiently mature and well suited to our research needs.

Based on these findings, the following critical features of effective learning environments for complex skill acquisition/generalizability provided the cornerstone for the Sherlock tutors:

Given the goal to accelerate instructionally the acquisition of complex and flexible problem solving skills:

- A pedagogical approach of **learning by doing** and **learning through reflection** in a situated, supported learning environment was implemented.

- Situated learning supplies the needed context for learners to execute tasks in an environment that reveals the **use** of their knowledge.
 - Supported learning enables a master-apprentice instructional relationship to form, whereby the master (or coach) can model, assist, and/or review problem solving performances as well as sequence learning activities that promote successive approximations of mature practice.
- Detailed cognitive models and authentic problem solving scenarios are essential as inputs to the instruction of complex problem solving skills.
- Cognitive models make the targeted expertise explicit, precise, and complete (e.g., tacit strategic knowledge is revealed).
 - Authentic scenarios give the instruction validity and vigor, with their realism, and they promote the culture of expert practice in the work environment.

Intelligent Computer Tutors Informed by Cognitive Models

Our goal in this section is to describe the resultant tutors -- known as Sherlock 1 and 2 -- but from an instructional perspective only. Our collaborators in this enterprise (and the exclusive software developers) at the Learning Research and Development Center, University of Pittsburgh, have written extensively on how the instructional functionalities are achieved in the software (Eggan & Lesgold, in press; Katz & Lesgold, in press; Katz & Lesgold, 1991; Katz, Lesgold, Eggan & Gordin, 1992; Lajoie & Lesgold, 1989; Lesgold, Lajoie, Bunzo, & Eggan, 1992).

Pedagogically, Sherlock functions as a coached, practice environment where students "learn by doing" electronic troubleshooting. They encounter high-difficulty fault isolation tasks which they pursue in a computer learning environment that is an extension of their real world work environment (situated learning). In the real world, technicians in this domain repair and maintain electronic subsystems and selected test equipment for the F15 aircraft. When a line replaceable unit (LRU or black box) is removed from the aircraft on the flightline because of a suspected malfunction, it is sent to technicians in the repair shops. Upon arrival it is attached via connecting cables (or similar apparatus) to a large piece of test equipment known as a test station. The LRU is then referred to as a Unit Under Test, or UUT. Figure 3 shows a top level diagrammatic view of the complete equipment system. A digitized picture of the test station is shown in Figure 4.

The tutor's troubleshooting scenarios are products of the cognitive task analysis (CTA). Similarly, the coaching is informed by the components of troubleshooting skill identified by the CTA data. In practice the distinctive characteristic of all Sherlock coaching is that it is under learner control, to be accessed **during** the problem solving process, as desired. This makes the

context of use of all knowledge apparent to the learner. Knowledge is not taught detached from its conditions of applicability. The benefits of accessing coaching are reinforced to the trainee in the post-problem reflective follow-up activities.

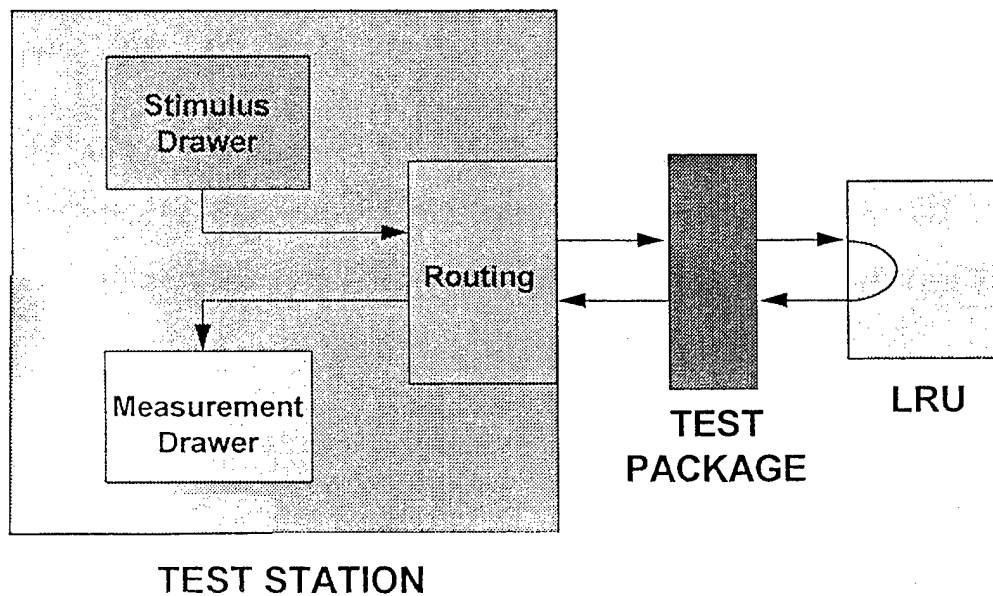


Figure 3. Top-Level Mental Model of Avionics Equipment System

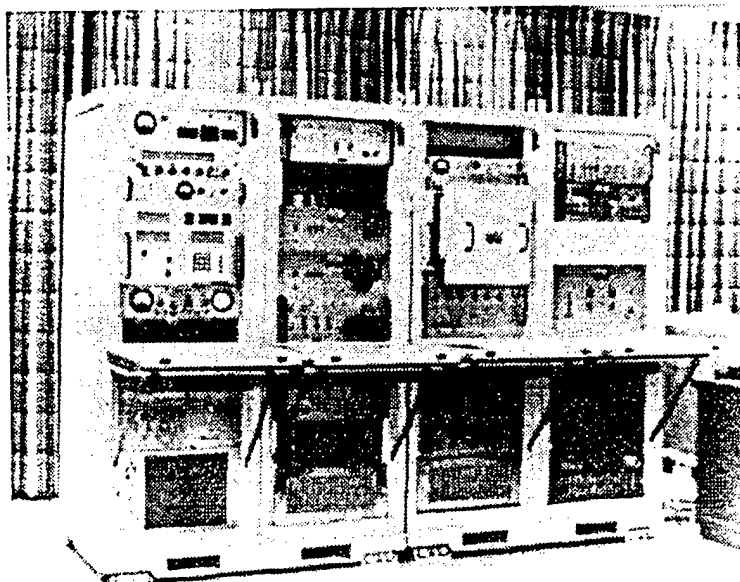


Figure 4. Actual Avionics Test Station

The coaching functionality is only one of the standard functionalities of intelligent tutoring systems such as Sherlock. A complete set of the tutor's functionalities is shown below (Polson & Richardson, 1988; Wenger, 1987).

- (1) A simulation of the equipment system (Device Model)
- (2) A simulation of the expert problem solver(s) (Expert Model)
- (3) Pedagogical knowledge (the Instructional Module), to include

- a graded series of problem solving scenarios
- coaching, to allow for supported learning
- decision rules for moving students through the scenarios
- methods for conducting post-problem reflective follow-up, i.e.,

interactions between coach and student about solution just generated

(4) A Student Model that tracks individual student progress, including problems solved, troubleshooting violations incurred, and so forth, and

- (5) An Interface that allows students to

- manipulate the front panel controls on the equipment and investigate the equipment system internally,
- access coaching, and
- participate in the reflective follow-up.

These instructional design features are supported by extensive empirical findings and theoretical models from cognitive science, as elaborated below.

Theory-Based Instruction

Results from studies of the acquisition of complex, practical skills (such as electronic troubleshooting) have demonstrated how cognitive theories of learning and performance can enable improved apprenticeship training (Gott, 1989). Improvements seem attributable to two general instructional advances, which roughly correspond to better content and better method. First, cognitive theoretical models provide detailed representations of expert task performance as the targets of instruction. In the details, the goals to which procedural knowledge applies and the strategic processes that are responsible for the organization, coherence, and general execution of the performance are clearly established. (See Figure 3.) In short, with cognitive models as input to instructional systems such as Sherlock 1 and 2, knowledge is directly tied to its uses in the world, and tacit knowledge (including goals, strategies, and assumptions) is made explicit for teaching. Content is thereby richer, more precise, and surrounded by context that establishes conditions of use.

Second, better method has been achieved through a union of modern formulations of skill acquisition and traditional apprenticeship training techniques, such as modeling and coaching (Collins, Brown, & Newman, 1989; Palincsar & Brown, 1984; Scardamalia & Bereiter, 1985; Schoenfeld, 1985). The common element in both is the notion of skill development as successive stages of increasingly mature performance. Hallmarks of apprenticeship training methods include situated learning, where students execute tasks and solve problems in an environment that reveals the various intended uses of their acquired knowledge; external support or scaffolding in the instruction in the form of ideal modeling of the performance, hints, reminders, explanations, or even missing pieces of knowledge to assist the apprentice's task execution; fading of external support as the apprentice's skill and autonomy build; and carefully sequenced learning activities that are both sensitive to changing student needs at different stages of skill acquisition and robust and diverse enough to foster integration and generalization of knowledge and skill (Collins, Brown, & Newman, 1989). Finally, to synthesize and reinforce the problem solving process, the solution steps are inspected, evaluated, and compared to examples of more advanced solutions at the end of each session. The Sherlock tutors were designed with these "better methods" as part of the instructional blueprint.

In sum, with better method, learning is assumed to occur through guided experience in instructional environments that provide progressive, explanation-based, and otherwise generally supported practice in the mechanics of solving problems. Intelligent tutoring systems informed by cognitive models enable the realization of this type of interactive and adaptive pedagogy.

With these major instructional design decisions having been implemented, we initiated the Sherlock 2 evaluation study to determine if an intelligent tutoring system that is informed by detailed cognitive models of troubleshooting performances is effective in both accelerating skill acquisition and fostering adaptive expertise. We conducted a controlled experiment that involved technicians at three geographically separated Air Force F15 flying wings.

METHOD

We evaluated trainees on a number of technical proficiency and experience indicators to establish matched experimental and control groups of apprentice technicians. We hypothesized that the experimental group would demonstrate an accelerated rate of skill acquisition compared to the control group trainees in moving toward the level of performance displayed by advanced technicians. These expectations were predicated on the following premise: given a learning environment that provides direct but coached problem solving experiences and one in which cognitive skill components and processes have been precisely identified as instructional targets, the acquisition of complex skills such as electronic troubleshooting can be speeded up. The expected accelerated acquisition would be attributable to better instructional content informed by cognitive models and better methods where learning by doing and learning by reflection are emphasized in a coached, computational practice environment.

We further hypothesized that the experimental group would demonstrate adaptiveness in their newly acquired troubleshooting skills compared to the control group when tested on a novel equipment system. These expectations were predicated on several premises. First, a learning environment where extensive practice is available to trainees would build up robust task schemas supported by conceptual support knowledge that explains the "reasons why" tasks are structured the way they are. Thus, knowledge structures would be both robust and flexible. Secondly, all coaching and post-session reflective feedback would provide general as well as task specific explanations to inject elasticity into system, procedural, and strategic knowledge components.

RESULTS

As predicted, posttest scores on several troubleshooting measures revealed large and statistically significant differences favoring the experimental group over the controls (Table 1). The pre- to posttest differences for both tests and groups are illustrated in Figure 5. Comparisons on both tests for all three groups (Master Group included) are illustrated in Figure 6.

Table 1. Posttest Measures of Troubleshooting Proficiency (Sherlock 2)

Sherlock 2 Posttests				
Group	N	Verbal Troubleshooting Test #1	Verbal Troubleshooting Test #2	Written Test on Troubleshooting
Novices				
Control	23			
		M	59 _a	58 _a
		SD	37	37
				75 _a
				14
Experimental	18			
		M	95 _b	91 _b
		SD	5	7
				87 _b
				12
Masters				
	13			
		M	85	86
		SD	12	11
				86
				11

Note. Means with different subscripts differ significantly at $p < .01$.

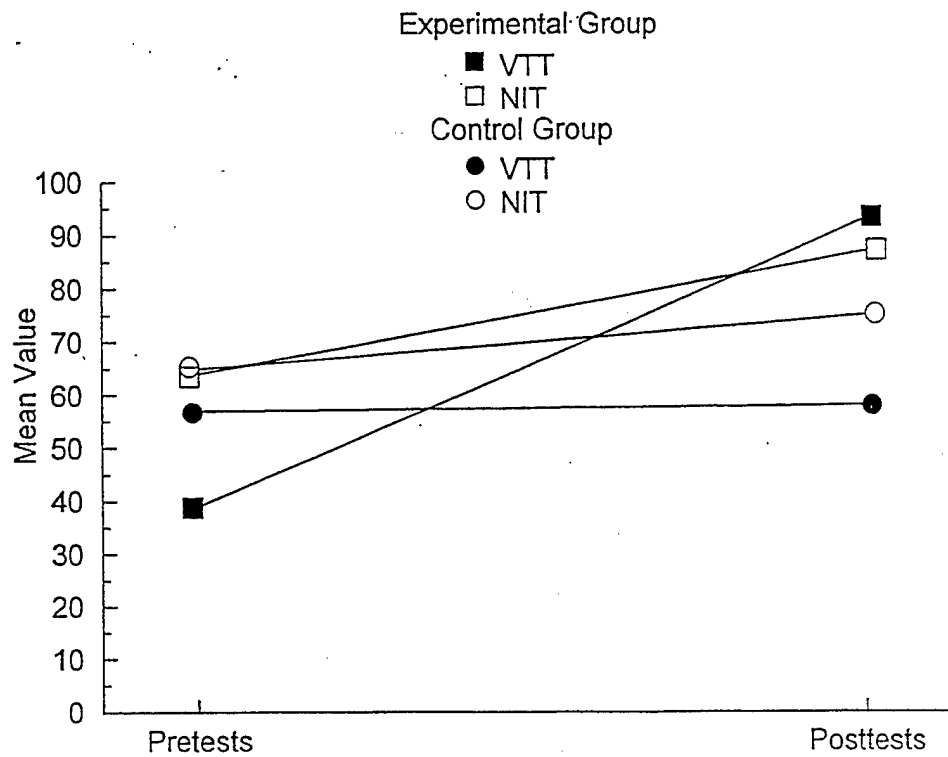


Figure 5. Pre to Posttest Changes: Sherlock 2 Tests

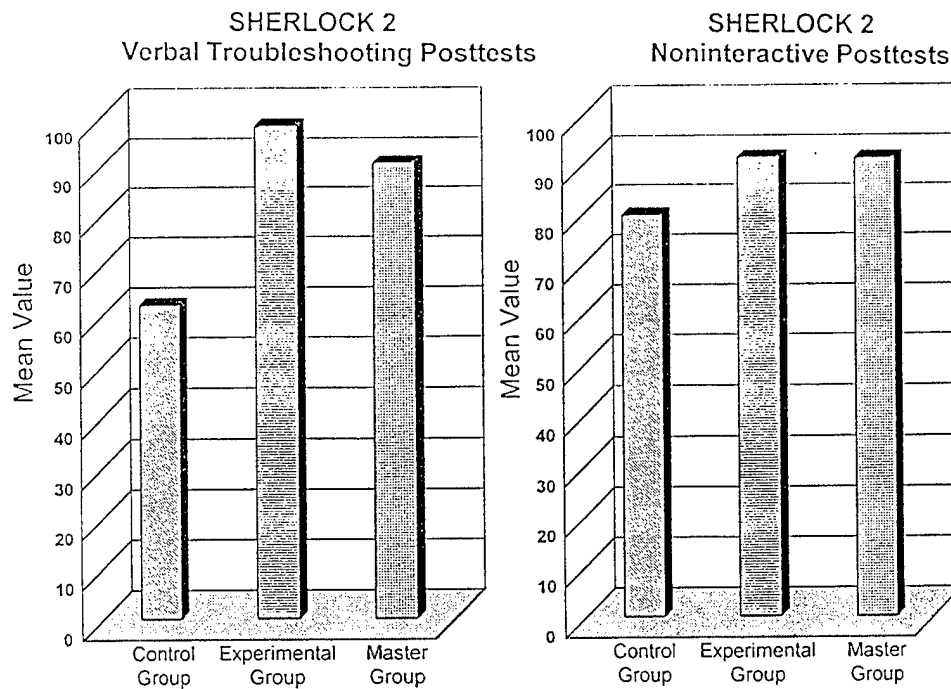


Figure 6. Sherlock 2 Posttest Scores Across Groups

Also, as predicted, the Frankenstation (novel equipment) posttest measures of transfer revealed statistically significant differences in favor of the experimental group over the controls (Table 2). The differences are illustrated in Figure 7.

Table 2. Posttest Measures of Transfer (Frankenstation)

Frankenstation Posttests			
Group	N	Verbal Troubleshooting Test	Written Troubleshooting Test
Novices			
Control	21		
<u>M</u>		55 _a	72 _a
<u>SD</u>		31	4
Experimental	17		
<u>M</u>		82 _b	80 _b
<u>SD</u>		23	10
Masters			
	12		
<u>M</u>		91	88
<u>SD</u>		22	12

Note. Means with different subscripts differ significantly at $p < .05$

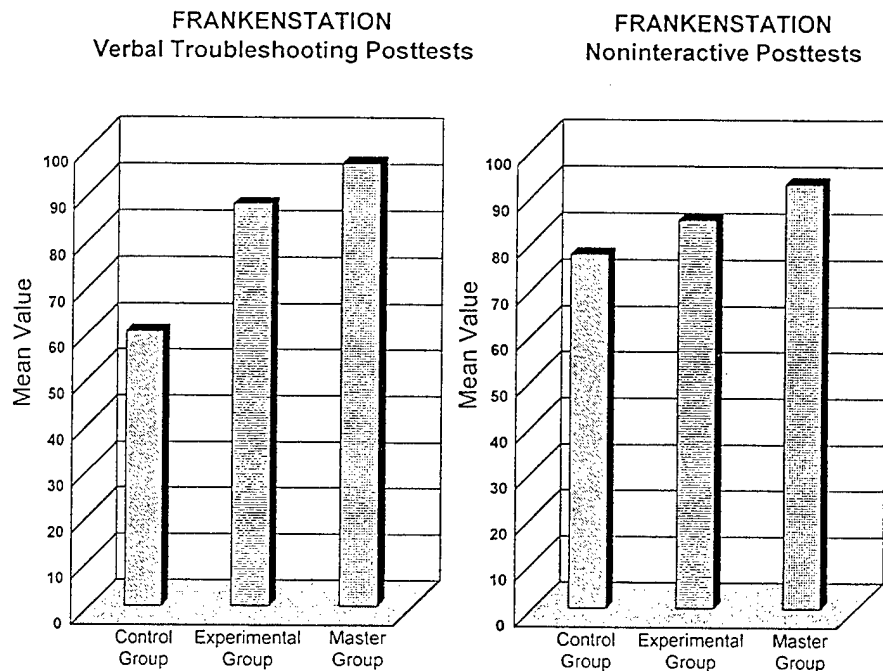


Figure 7. Frankenstation Test Scores Across Groups

The effect size for each of the posttest measures is shown in Table 3. As a basis for comparison, the average effect size for new science and math curriculum in U.S. schools is reported to be .3 standard deviation (Bloom, 1984).

Table 3. Effect Size for Posttest Measures

Measure	Group						Effect Size
	Control			Experimental			
	N	M	SD	N	M	SD	
VTT 3 (Sherlock)	23	59	37	18	95	5	1.27 SDS
VTT 4 (Sherlock)	23	58	37	18	91	7	1.17 SDS
NIT (Sherlock)	23	75	14	18	87	12	.87 SDS
VTT (Frankenstation)	21	55	31	17	82	23	.96 SDS
NIT (Frankenstation)	21	72	11	17	80	10	.76 SDS

DISCUSSION

Experience Equivalence Analyses

One way to express the practical impact of Sherlock 2 is to translate posttest differences between the Tutored and Control Groups into "equivalent man-months experience." That is, a practical interpretation of the experimental outcome would be to determine how many months of on-the-job experience one would need to move from the pretest level of performance to the posttest level.

In the analysis of Sherlock 1 evaluation data, this same approach was used in separate analyses of Sherlock test scores (Gott, 1989; Nichols, Pokorny, Jones, Gott, & Alley, in preparation). The impact of the tutor in terms of experience gained was between 48 and 52 months for the two types of tests, meaning the tutored group gained an amount of experience equivalent to approximately 50 months. The added experience "value" could be attributed to the hours spent working problems on Sherlock 1, which was approximately 20 hours spread over 12 working days. In these practical terms then, the cost of Sherlock 1 can be characterized as approximately 20 hours of tutoring per airman, and the benefit as approximately 50 months of added experience per airman.

In attempting to replicate this analysis with Sherlock 2 data, some weighty differences in test composition rendered the replication implausible. The Sherlock 1 tests were broader in scope than the analog tests for Sherlock 2. This occurred because with Sherlock 2, the most difficult troubleshooting scenarios had to be deleted when the tutor curriculum was truncated at Site 1 (Langley AFB VA) in order to meet schedule commitments. Consequently, the Sherlock 2 Verbal Troubleshooting Tests were in effect made easier, and posttest scores for Tutored airmen were quite high, reflecting a ceiling effect for those tests. Frankenstation scores were, however, not affected by the shortening of Sherlock tests, since that novel equipment system was constructed solely to measure anticipated transfer effects.

The Frankenstation scores were thus deemed more meaningful for purposes of the experience equivalence analysis. Frankenstation verbal test data have greater variability for tutored airmen than Sherlock 2 verbal data: mean of 81.77, SD 23.06, minimum of 31.73 and maximum of 122. There is also a strong relationship between manmonths experience and troubleshooting performance on this test ($r = .75$, $N = 17$, one-tailed significance $p < .001$). The impact of Sherlock in terms of fostering generalizable troubleshooting skills (i.e., skills transferrable to novel equipment) was 65.6 months, or about 5.5 years. The value of 5.5 years is fairly consistent with the Sherlock 1 experience equivalence results.

How has Sherlock 2 been able to achieve this level of effectiveness? Although our study did not include conditions where instructional features were manipulated to gauge the power of individual instructional attributes, past research in the training and transfer of complex problem solving skills lends support to certain speculative explanations regarding Sherlock's effectiveness. The support centers around (a) cognitive models as input to instruction; (b) situated learning in a constructivist instructional environment, and (c) the sociology of a learning system. In terms of Sherlock's effectiveness in fostering generalizable skills, we offer speculative explanations that concern (a) the quality of the performer's knowledge representations in the source domain; (b) the availability of a top level plan or goal structure; and (c) multilevel structures of varied content that have become highly practiced. We now discuss each of these dimensions in turn.

Cognitive Models as Input to Instruction

One way to speculate about Sherlock's effectiveness is to contrast it with other troubleshooting training systems, which have generally demonstrated meager learning effects. The nature and derivation of instructional content is the vital point of focus. With Sherlock, explicit, multicomponent cognitive models of troubleshooting performance derived from Master technicians provided the input; whereas, other instructional systems typically focused on one, or at best, two components of troubleshooting expertise. Further, the content of other systems has rarely come directly from expert performers via knowledge elicitation techniques.

Training that focuses exclusively on procedures (as a targeted troubleshooting component) is limited in numerous ways. The content usually takes the form of hundreds of specific production rules as curriculum, which can, except for the simplest of systems, cause cognitive overload. Also, the rules are not usually organized into meaningful chunks to facilitate learning. Instead, the procedures (rules) are presented instructionally as a flat string of actions. Flat strings of procedures do not correspond to the internalized procedural representations of experts. Their procedural knowledge is organized as hierarchical, goal-directed structures. Instruction, therefore, is not informed by the expertise that is the ultimate performance to be acquired. Further, procedures, as independent, situation-specific rules, have limited generalizability and represent only a partial model of performance, since the explicit teaching of device and strategic knowledge components is typically ignored in procedures-based training.

Similarly, in device knowledge-based instruction, facts and formal laws are too often taught as "detached" pieces of declarative knowledge. The content is not tied to its conditions of use in the real world, that is, when and how it is useful in problem solving contexts. As a result, students have no direct manipulation experiences with the facts and formal laws about a system or device. It is manipulation experiences that promote the building of cause and effect relationships and the capability to **infer** procedures (Hollan, Hutchins, McCandless, Rosenstein, & Weitzman, 1987). Such inferencing skills are extremely important in complex, ill-structured domains where it is virtually impossible to prespecify every series of actions to take. Novel situations always arise, and the problem solver then must use **device** knowledge to **construct** a series of **procedural steps**.

Situated Learning in a Constructivist Environment

Following Dewey and in accord with the current constructivist movement (Perkins, 1991), the general principle of **learning by doing** is the touchstone of the Sherlock instructional design. The tutor is an extension of the trainee's actual work environment. Authentic fault isolation problems are selected and presented to students as holistic scenarios to solve as they actively construct their understanding of the equipment and of the troubleshooting task. Working in Sherlock is like doing one's job in the real world -- objects in the environment are acted upon to achieve certain goals. There are, however, several nontrivial bonuses in Sherlock that do not exist in the real world.

First, in the actual shop environment, trainees must learn about test station troubleshooting as opportunities present themselves. Unfortunately, the opportunities are infrequent, because the frequency of station malfunction is relatively low. However, being able to successfully troubleshoot the test station is a critical skill for this specialty, commanding high training emphasis. Further, since malfunctions occur essentially at random, learning is driven by whatever breaks. Instruction cannot be sequenced in the manner just described where movement through upwardly compatible models of understanding and performance can be fostered. Finally, it is possible in Sherlock to time-compress the routine activities that may take an inordinate

amount of time in the real world so that valuable instructional time is devoted to the challenging part of the task.

The second bonus is that the Sherlock environment is forgiving; mistakes can be made without dire consequences, plus, expert coaching is always available as scaffolding, when needed. Scaffolding in a learning environment supports trainees as they try to make sense of the domain -- with hints, explanations, even missing pieces of knowledge. In Sherlock 2, the scaffolding appears as coaching during problem solving, and additional support is provided in reflective follow-up activities.

Coaching

The coaching available to students while solving problems corresponds to the components of skill (system, procedural, and strategic knowledge) revealed by the cognitive task analysis. "How It (the system) Works," "How to Test (the system with procedures and strategy)", and "How to Trace (the schematics that define the system)" comprise the coaching that scaffolds performance at multiple levels of specificity. The trainee's objective is to learn from the coaching but ultimately grow independent of it, just as traditionally, the apprentice gains independence from the master.

Even without extensive use of coaching, the student can explore and investigate the equipment as phenomenaria (Perkins, 1991) in the simulated environment as he attempts to make sense of the task and the equipment system. What seems crucial to the learning process in Sherlock is the presence of mystery:¹ there is always a problem to be solved. Not only is this a motivational hook, but the situated nature of the fault diagnosis allows the student to experience the conditions under which procedures are being used and how device and strategic knowledge empower the problem solving performance. In short, the content and task are meaningful.

One important component of the constructivist approach that is apparent in the coaching of Sherlock 2 is the intelligent hyperdisplay. **How it Works** and **How to Test** coaching are accompanied by abstracted schematic diagrams to help illustrate the advice the coach is providing. The frequent use of illustrations in the tutor is in response to the heavy role played by equipment mental models in expert performances. Pedagogically, the goal is to expose learners to the illustrated models so that they begin to incorporate them into their own problem solving, envisioning the system in their mind's eye as they engage in diagnosis. The structure of the diagram reflects the expert representation of the circuitry involved in carrying out the function that failed, as revealed in the detailed cognitive task analysis. What is displayed is approximately what a trainee would want to know at that time, but every display component is "hot" and can be used as a portal to more detail or explanation. The part of the system on which the expert would

¹ We are indebted to Allan Collins of BBN and Northwestern University for sharing with us his beliefs about the importance of incorporating a sense of mystery in a learning activity, i.e., some "puzzle" to figure out to keep the learner actively engaged.

be focusing at a given point in the problem solution process is allocated the most space in the diagram and presented in the most detail.

From our earliest ethnographic interventions into this domain of electronic troubleshooting, we have been impressed with the quality of the performers' mental models and the central role they play in the quality of their troubleshooting. Two contrasting models are shown in Figures 8 and 9 to illustrate the point. The model revealed by the weak performer (weak in terms of troubleshooting proficiency) shows that his internal representation of the F15 antenna system to be based on physical structure and sparsely detailed. By comparison, the strong performer reveals a richly detailed, functional representation -- one that gives him a powerful blueprint in his mind's eye to guide his troubleshooting. His functional model shows that his understanding of how the system works has penetrated the opaque surface features of the antenna system. Both these sketches were produced in response to the question: "When you work on the antenna, how do you think about it? How would you draw it on paper?"

Weak Performer:
physical decomposition

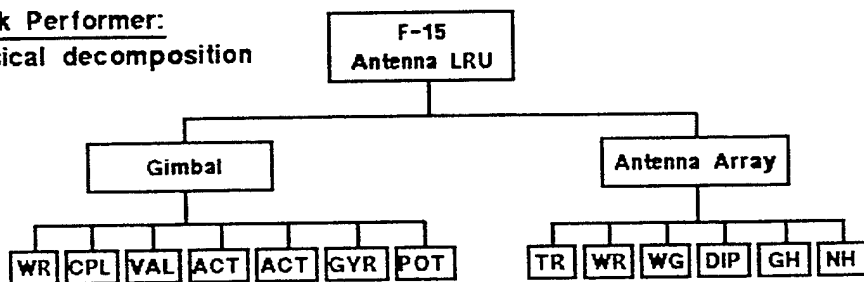


Figure 8. Example of Weak Performer's Mental Model

Strong Performer:
functional decomposition

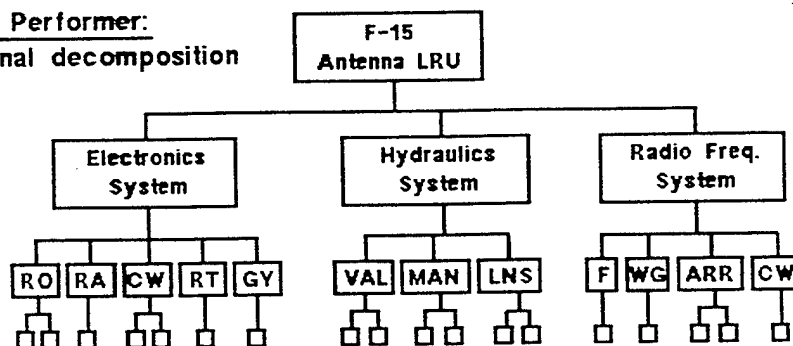


Figure 9. Example of Strong Performer's Mental Model

Reflective Follow-up (RFU) Activities

In the Sherlock 2 RFU, students engage in four activities designed to foster learning through **reflection**: (a) they view a replay of the solution steps they just executed; (b) they see their solution juxtaposed to an exemplar Master solution; (c) they can view a more elaborated replay of a Master solution; and (d) they are asked to diagnose their own solution trace using principles of good troubleshooting espoused by Masters in this domain. Their self-diagnosis is then compared to the computer coach's diagnosis, for validation. The RFU culminates the problem solving session, and, we believe, provides a considerable value-added element to the Sherlock 1 system where there was no post-performance review capability.

By enhancing the earlier Sherlock 1 **learning-by-doing pedagogy** with a **learning-by-reflection component**, we believe the instructional impact has been heightened. As Collins and Brown (1988) and Collins, Brown, & Newman (1989) have argued, learning through reflection in a computational environment such as Sherlock achieves the following objectives:

(1) The student's solution trace becomes a useful object of study, especially since the computer can represent the **process** of the solution and thereby externalize decisions for interpretation from a variety of perspectives.

(2) By having access to an exemplar solution from a Master (including the Master's normally tacit reasons for each action), the student can observe and even discover expert strategies and reasoning that subsequently can improve the trainee's own solution.

(3) After viewing a number of problem-specific traces in the RFU, the student can derive abstractions from the patterns of actions and underlying reasons, e.g., a generic goal structure.

(4) By treating each of his traces as useful objects of study, the student can come to view learning as an "incrementally staged process" that happens over time, not all at once (Collins, Brown, & Newman, 1989). Further, the self-diagnosis activity provides the student with concrete benchmarks of his own progress along the skill continuum.

(5) The self diagnosis that occurs can then become internalized by the student as a form of self-correction, self-monitoring capability. These are metacognitive skills the trainee may not have possessed before the Sherlock experience.

Our confidence in the power of the RFU was bolstered from personal observations made during the field trials. It was not uncommon for a trainee to remark (with undisguised delight), "Oh, that's cool the way he (the Master) ruled out the Test Package with just one measurement. I've been making two or more tests to do the same thing. I'm going to do it his way next time." A sampling of other post-performance comments include the following:

"Some of this stuff is already helping me in the

shop. I find I'm being more efficient in a number of ways."

"I have a lot more confidence now about when to take an ohms measurement and when to test for voltage. I was always confused about that before."

We also witnessed and received feedback from trainees who talked productively about the tutor with Master technicians in the shop. This phenomenon provides a natural transition to the sociology surrounding Sherlock and how the sociological aspects of an instructional system can clearly contribute to its overall effectiveness. We treat that topic next.

Sociology of the Learning Environment

Above all else, instruction must be viewed as valid to trainees. It must serve their needs, profit them directly. "Drawing students into a culture of expert practice in cognitive domains involves teaching them to think like experts" (Collins, Brown, & Newman, 1989, p. 488). In effect, Sherlock seeks to do just that -- draw students into a culture of expert practice, enable them to reach the mature levels of proficiency they observe being practiced by the masters who are the acknowledged leaders in the shop.

On a daily basis during the field trials we observed the tutored subjects make strides in becoming a part of the community of expert practice. One could argue they were rediscovering learning along with us. They shared with us conversations they had with their Team Leaders when malfunction problems arose on the actual equipment that were covered in Sherlock. They sometimes consulted the acknowledged masters in the shop (after a tutoring session) when they needed and wanted more elaboration about a Sherlock scenario than was available in the tutor. There were also occasions during tutoring sessions when trainees would want to get an opinion from one of the shop 7-levels because they thought the tutor's interpretation or suggestion was too narrow or incomplete.

For some of the more experienced apprentice subjects, they sometimes opted to use the device simulation in the tutor as a way to **discover** the equipment's functionality and observe general cause and effect relationships. They did this by making additional measurements during a Sherlock scenario (i.e., more than needed to solve the problem) just because they wanted to verify or expand their device knowledge. They explained that when they first started working on another test station, e.g., the Electronic Warfare Station, they would do just that -- take a lot of measurements as a way of figuring out how the system worked.

These observations provide one measure of Sherlock's effectiveness in socializing apprentices into the expert culture. On the Tutor Report Card, trainees also reported increased confidence in performing the hardest tasks in their job. The argument can be made then that

Sherlock was viewed as valid instruction by trainees. They were learning to perform tasks that they recognized as having value in the shop, in their culture. Moreover, the acquiring of skill and knowledge from Sherlock was enabling them to be more conversant with shop Masters about the domain. In short, their status in their culture was on the rise and it is reasonable to assume they attributed some of that ascendance to Sherlock.

Quality of Knowledge Representations -- Runnable Mental Models

To be effective in today's technology-dense and highly dynamic work environments, the performer must be a good transferer of knowledge and skill. There is a growing body of evidence that suggests that adaptiveness in generalizing one's knowledge -- particularly in the context of complex problem solving tasks -- is strongly influenced by the **quality** of the performer's knowledge representations (Brown & Burton, 1986; Brown & deKleer, 1985; Chi & Bassok, 1989; Chi, Feltovich, & Glaser, 1981; Clancey, 1986; Judd, 1908; Katona, 1940; Kieras, 1982; Kieras & Bovair, 1984; Wertheimer, 1945; White & Frederiksen, 1986, 1987). Adaptive expertise has in fact been posited as an advanced level of problem-solving performance that is characterized by principled representations of knowledge and skill as opposed to representations dominated by surface features (Hatano & Inagaki, 1984). For example, the procedural skills of adaptive experts are supported by reasons and other conceptual support knowledge that explain the steps of the procedure including the conditions of applicability.

When we first began to investigate the process of transfer in this domain, we conducted interviews with a broad range of technicians (Gott, Hall, Pokorny, Dibble, & Glaser, 1992). The results of these early interviews suggested some interesting transfer-related differences across levels of technical proficiency. Skilled performers appeared to use existing schemata from their source domain to build representations in the target domain. More specifically, their stated expectations revealed existing schemata of three general types: first, (mental) device models of the avionics equipment, second, general models of the task, in the form of a troubleshooting approach (or goal structure); and third, well-organized procedural knowledge that appeared to be adaptive to the conditions imposed by more and less complex equipment systems.

With respect to system knowledge, the better performers talked about the equipment in terms of an abstracted, integrated device representation based on system functionality: "A test station is a test station. The same general functions are performed by every station -- a stimulus is generated, routed, and measured. The electronic test process is the same. Most differences are in the physical aspects of the equipment." Even before procedural steps are taken, skilled performers appear to represent a problem in terms of some meaningful structure or functionality whereas less-skilled performers organize their representations around surface features and literal descriptive aspects (Chi, Feltovich, & Glaser, 1981; Glaser, 1988; Gott, Hall, Pokorny, Dibble, & Glaser, 1992). This type of functional qualitative analysis often involves representing a problem in terms of a **"runnable mental model"** which specifies the main causal connections of the

components used in a test situation. Running an internalized model can be thought of as invoking an explanatory theory to use in instantiating a given problem situation.

We now treat in more depth the runnable models that the tutored airmen generalized to the novel system, Frankenstation. On the Frankenstation verbal troubleshooting test, the high performers in the tutored group took one of two approaches in solving the problem. Both approaches strongly suggested that a device model from Sherlock 2 was directing their actions. In terms of their approaches to Frankenstation, they either (a) used schematics to instantiate their mental model of general equipment functionality (see Figure 3) and then took measurements guided by the instantiated model to localize the fail and solve the problem or (b) they used schematics and/or the computer program listing to instantiate the functional mental model, then selectively used self-tests to localize the fail, and finally made the necessary measurements to solve the problem. In general, both groups proceeded in a systematic way, using their deep functional representations from Sherlock to guide their actions on Frankenstation. The inferences we draw about the quality of their representations become more credible once the behavior of the untutored novices is revealed.

The most telling characteristic of the traces of the untutored airmen may be the manner in which they used self-tests. Many of them initiated the self-test process from the beginning, where the Digital Multimeter (DMM) was the first device tested. The DMM was not even being used in this particular Frankenstation test. Therefore, they did not appear to be instantiating a functional mental model by identifying a measurement device to be investigated via a self test. It was not until the seventh self-test where an active device was targeted. Had the control group airmen been instantiating a strong, internalized equipment model, they would not have been wasting their time investigating inactive station resources such as the DMM.

Presence of a Top-Level Goal Structure

Our early work to investigate transfer in this domain revealed that skilled performers not only **transported** abstract device schemata to a new target domain (as we just discussed), but they also generalized their well-practiced planning or goal structure schemata as well. In Sherlock 2, emphasis was placed on logically investigating the equipment system in a sequence that was extracted from, supported by, and then, in the tutor, modeled by master technicians. The masters' reasons for the sequence were made explicit so that trainees did not mindlessly learn what to investigate first, second, third, etc. Reasons centered around a cost-benefit ratio process whereby they estimated the information value of an action versus the cost, such as likelihood of fail, time, physical effort, cognitive effort, and so forth.

The tutored airmen made significantly fewer deviations from a logical sequence than did the control group. Interestingly, the Master group also recorded numerous violations of the logical sequence. Upon closer examination, however, what we see in Master solutions is the expected use of short cuts that are perfectly acceptable in the real world shop environment;

however, when teaching apprentices, the goal is to ensure that they learn system knowledge, not procedural short-cuts that would result in their avoidance of critical "how it works" knowledge, which is the centerpiece of skilled troubleshooting.

Multilevel Structures to Transfer

In a technical domain such as electronics problem solving, we expect transfer (as learning) to occur on many levels, such that a range of specific-to-general activities of either **transporting** or **transposing** prior knowledge takes place. For example, procedural skills such as the execution of standard electronic tests (e.g., ohms and voltage measurements) represent one of multiple types of knowledge that must be effectively coordinated to produce optimal levels of reasoning in a complex technical domain (Gott, 1989). The procedural steps of such standard electronic tests may be fairly constant across equipment systems. As a result, this rule-based knowledge may be transportable intact across job domains in a manner that suggests a common (or identical) elements form of transfer (Thorndike, 1903, 1906; Thorndike & Woodworth, 1901). Thus, transferring compact procedural subroutines such as the steps of an ohms test would represent an instance of specific procedural transfer (or self-transfer). By contrast, the transfer of the type of abstract representations we have just discussed, such as the goal structures and device schemata that are used to deploy and guide specific problem-solving steps, would constitute transfer of a more general nature

Even with the availability of compact procedural subroutines from the source domain, we observed adaptiveness from tutored airmen when conditions dictated. For example, the experimental subjects who initially tried to measure computer bit data with a digital multimeter typically recovered from that procedural error, realized they needed a different piece of test equipment, and made the appropriate adjustment. Also, it is important to note that the Frankenstation problem involved a square wave signal which requires the use of an oscilloscope. Because the Segment 2 problems in Sherlock (i.e., the stimulus problems where oscilloscope use would be required) had to be aborted, tutored subjects had received no training with this type of test equipment on the tutor. In short they had little if any experience with such procedures in their source domain.

SYNTHESIS AND SUMMATION

In considering the rediscovery of learning in complex real world tasks, I have posed two major premises as foundation. First, skill acquisition in practical domains depends upon purposeful learning experiences where knowledge connects with its uses in the world. Secondly, cognitive modeling makes explicit the forms and utilities of knowledge that may otherwise go unobserved, untaught, and therefore unlearned. These premises have come to the forefront of practical skills training as a result of several significant educational and technological trends.

Educational systems at all levels appear to have gradually weakened the ties between the knowledge/skills (that are the province of formal schooling) and their uses in the world. A renewed interest in apprenticeship instruction and related empirical work in learning suggest that the pedagogical principles that characterize classic apprenticeship methods can help to remedy this discontinuity.

Further, as rapidly advancing technologies increase the number of mental (vs. manual) tasks, critical elements of the expert's performance are correspondingly more likely to become unobservable to the apprentice. Mental processes and features of knowledge often remain tacit, that is, unarticulated and therefore unknowable. Cognitive models of real-world task performance make the elements and processes of modern-day expertise explicit, observable and therefore potentially learnable.

In practical domains, the modeled expertise has revealed the multiple sources and levels of knowledge that experts bring to bear on the types of complex, ill-structured problems that are commonly encountered in modern-day workplaces. A major finding concerns the expert's capability to engage in adaptive, opportunistic reasoning that involves the coordination of three major sources of knowledge: procedural, device (or system), and strategic control knowledge. For each type of knowledge, the expert has access to elaborate abstraction hierarchies that range from specific knowledge instantiations to abstractly stated principles.

A second major conclusion, which is supported by a growing body of empirical evidence, is that skill acquisition occurs through successive approximations of the targeted expertise. The progression is characterized by movement from partial to more complete (and thus complex) (a) **understandings** of domain phenomena, (b) **procedural** subroutines, and (c) **strategic control structures**, including goals, plans, and decision factors. The apprentice achieves various levels of incomplete knowledge and capability on the way to mastery. Principles and methods from classic apprenticeship training are proving useful in fostering this progression. **Situated, supported,** and carefully devised and **sequenced** learning experiences have been shown to foster development.

Concordant with these premises, I report results from the recent controlled experiment conducted to evaluate a U. S. Air Force avionics troubleshooting tutor (Sherlock 2). Outcomes of **situated and supported learning** in this tutor show dramatic gains in fault isolation proficiency, as well as in skill generality by apprentice technicians. With the Sherlock (computer tutor) system, authentic fault isolation scenarios are presented in a computer environment that is an extension of the airman's real work environment. Learning is meaningful because the trainee works on problems he sees the acknowledged experts in his culture confront every day. He observes his own incremental growth and eventual movement into this same community of expert practice. He even gains confidence he can transfer what he knows to novel equipment systems. A victory for successful learning in the modern workplace can be declared.

RECOMMENDATIONS

Three advances in cognitive science have been central to the realization of this instructional system: (a) **cognitive performance models**, which Newell and Simon (1972) said over 20 years ago must be given precedence in research (over studies of learning processes) so that a more complete and successful theory of learning and development can ultimately emerge; (b) a **cognitive task analysis methodology** robust enough to codify the expert performance models, as well as the successive approximations of expertise encountered along the way; and (c) **the "mental experiments"** that experts run to simulate explanatory models ("in the mind's eye") during diagnostic and procedural tasks. The confluence of these and other advances have taken us a step forward in fostering, even rediscovering (if you wish), how learning in the real world can be engaging and profitable as never before.

REFERENCES

- Anderson, J. (1984). Learning to program in LISP. *Cognitive Science*, 8(2), pp. 87-129.
- Anderson, J.R., Boyle, C.F., Farrell, R.G., & Reiser, B.J. (1984). Cognitive principles in the design of computer tutors. Proceedings of the Sixth Cognitive Science Society Conference (Boulder, CO), pp.2-8.
- Bloom, B.S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, 13, 4-16.
- Brown, A.L. (1978). Knowing when, where, and how to remember: A problem of metacognition. In R.Glaser (Ed.), *Advances in instructional psychology* (Vol. 1, pp. 77-165). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Brown, A.L., Bransford, J.D., Ferrara, R.A. & Campione, J.C. (1983). Learning, remembering, and understanding. In J.H. Flavell & E.M. Markman (Eds.), *Handbook of child psychology* (Vol. 3, pp. 77-166). New York: Wiley.
- Brown, J.S., & Burton, R.R. (1986). Reactive learning environments for teaching electronic troubleshooting. In W.B. Rouse (Ed.), *Advances in man-machine systems research* (pp. 65-98). Greenwich, CT: JAI Press.
- Brown, J.S., & deKleer, J. (1985). A qualitative physics based upon confluences. In D.G. Bobrow (Ed.), *Qualitative reasoning about physical systems*. Cambridge, MA: MIT Press.
- Card, S., Moran, T.P., & Newell, A. (1983). *The psychology of human-computer interaction*. Hillsdale, NJ: Erlbaum.
- Chi, M.T.H., Feltovich, P.J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5, 121-152.
- Chi, M.T.H., & Bassok, M. (1989). Learning from examples via self-explanations. In L.B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 251-282). Hillsdale, NJ: Erlbaum.
- Clancey, W.J. (1986). Qualitative student models. *Annual Review of Computer Science*, 1, 381-450.
- Collins, A., Brown, J.S., & Newman, S. (1989). Cognitive apprenticeship: Teaching the craft of reading, writing, and mathematics. In L.B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser*. (pp.453-493). Hillsdale, NJ: Lawrence Erlbaum.
- Collins, A., & Brown, J.S. (1988). The computer as a tool for learning through reflection. In H. Mandl & A. Lesgold (Eds.), *Learning issues for intelligent tutoring systems*. (pp. 1-18). New York: Springer-Verlag.

- deKleer, J. (1984). How circuits work. *Artificial Intelligence*, 24, 205-280.
- di Sessa, A.A. (1982). Unlearning Aristotelian physics: A study of Knowledge-based learning. *Cognitive Science*, 6, 37-75.
- Eggan, G., & Lesgold, A. (in press). Modeling requirements for intelligent training systems. In S. Dijkstra, H. P. M. Krammer, & J. J. G. van Merrienboer (Eds.), *Instructional-design models in computer-based learning environments*. Heidelberg: Springer-Verlag.
- Ericsson, K.A. & Simon, H.A. (1993). *Verbal reports as data (rev.ed.)* Cambridge MA: MIT Press.
- Frederiksen, N. (1984). The real test bias: Influences of testing on teaching and learning. *American Psychologist* 39(3), 193-202.
- Gitomer, D.H. (1984). *A cognitive analysis of a complex troubleshooting task*. Unpublished doctoral dissertation, University of Pittsburgh, Pittsburgh, PA.
- Glaser, R., Lesgold, A., Lajoie, S., Eastman, R., Greenberg, L., Logan, D., Magone, M., Weiner, A., Wolf, R., & Yengo, L. (1985). Cognitive task analysis to enhance technical skills training and assessment. (Contract No. F41689-83-C-0029). Brooks AFB, TX: Armstrong Laboratory, Human Resources Directorate.
- Glaser, R. (1988, January). *Expert knowledge and processes of thinking*. Paper presented at the American Association for the Advancement of Science Meeting, San Francisco, CA.
- Gott, S.P. (1987). *Assessing technical expertise in today's work environments*. Proceedings of the 1987 ETS Invitational Conference, Princeton, NJ: Educational Testing Service.
- Gott, S. P. (1989). Apprenticeship instruction for real world cognitive tasks. In E. Z. Rothkopf (Ed.), *Review of Research in Education*. Vol. XV. Washington, DC: AERA.
- Gott, S.P., Bennett, W., & Gillet, A. (1986). Models of technical competence for intelligent tutoring systems. *Journal of Computer-Based Instruction*, 13(2), 43-46.
- Gott, S. P., Hall, E. P., Pokorny, R. A., Dibble, E., & Glaser, R. (1992). A naturalistic study of transfer: Adaptive expertise in technical domains. In D. K. Detterman & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction*. (pp 258-288). Norwood NJ: Ablex Publishing Corp.
- Gott, S.P., & Pokorny, R. (1987). *The training of experts for high-tech work environments*, Proceedings of the Ninth Interservice/Industry Training Systems Conference (pp. 184-190.) Washington, DC.
- Greeno, J.G. (1978). A study of problem solving. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. 1, pp. 13-75). Hillsdale, NJ: Erlbaum.

- Hall, E.P., Gott, S.P., & Pokorny, R.A. (in press). *A procedural guide to cognitive task analysis: The PARI methodology*. (AL/HRM Technical Report). Brooks AFB, TX: Human Resources Directorate, Armstrong Laboratory.
- Hatano, G., & Inagaki, K. (1984). *Two courses of expertise*. In the annual report of the Research and Clinical Center for Child Development. Sapporo, Japan: Hokkaido University.
- Hollan, J.D., Hutchins, E.L., McCandless T.P., Rosenstein, M., & Weitzman, L. (1987). Graphic interfaces for simulation. In W.B. Rouse (Ed.), *Advances in man-machine systems research*. (Vol. 3, pp. 129-163). Greenwich, CT: JAI Press.
- Judd, C.H. (1908). The relation of special training and general intelligence. *Educational Review*, 36, 28-42.
- Katona, G. (1940). *Organizing and memorizing*. New York: Columbia University Press.
- Katz, S. & Lesgold, A. (1991). *Modeling the student in Sherlock II*. In J. Kay & A. Quilici (Eds.), *Proceedings of the IJCAI-91 Workshop W.4: Agent modelling for intelligent interaction* (pp. 93-127). Sydney, Australia.
- Katz, S., & Lesgold, A. (in press). The role of the tutor in computer-based collaborative learning situations. In S. Lajoie & S. Derry (Eds.), *Computers as cognitive tools*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Katz, S., Lesgold, A., Eggan, G., & Gordin, M. (1992). Self-adjusting Curriculum Planning in Sherlock II. *Lecture Notes in Computer Science: Proceedings of the Fourth International Conference on Computers in Learning (ICCAL '92)*. Berlin: Springer Verlag.
- Kieras, D.E. (1982). *What people know about electronic devices: A descriptive study* (Technical Report No. 12 UARZ/DP/TR-82/ONR- 12). University of Arizona, Department of Psychology.
- Kieras, D.E. (1987). *What mental model should be taught: Choosing instructional content for complex engineered systems* (Tech. Rep. No. 24 TR-87/ONR-24). University of Michigan.
- Kieras, D.E. & Bovair, S. (1984). The role of mental model in learning to operate a device. *Cognitive Science*, 8, 255-273.
- Lajoie, S., & Lesgold, A. (1989). Apprenticeship training in the workplace: Computer coached practice environment as a new form of apprenticeship. *Machine-Mediated Learning*, 3, 7-28.
- Larkin, J.H., McDermott, J., Simon, D.P., & Simon, H.A. (1980). Expert and novice performance in solving physics problems. *Science*, 208, 1335-1342.
- Lesgold, A. M., Lajoie, S. P., Bunzo, M., & Eggan, G. (1992). SHERLOCK: A coached practice environment for an electronics troubleshooting job. In J. Larkin & R. Chabay (Eds.), *Computer*

assisted instruction and intelligent tutoring systems: Shared issues and complementary approaches (pp.201-238). Hillsdale, NJ: Lawrence Erlbaum Associates.

- McCloskey, M., Caramazza, A., & Green, B. (1980). Curvilinear motion in the absence of external forces: Naive beliefs about the motion of objects. *Science*, 210, 1139-1141.
- Means, B., & Gott, S.P. (1988). Cognitive task analysis as a basis for tutor development: Articulating abstract knowledge representations. In J. Psotka, D. Massey, and S. Mutter (Eds). *Intelligent tutoring systems: Lessons learned*. Hillsdale, NJ: Erlbaum.
- Mitchell, C. (September 28, 1987). Corporate classes: Firms broaden scope of their education programs. *The Wall Street Journal*, p.27.
- Newell, A., & Simon, H.A. (1972). *Human problem solving*. New York: Prentice-Hall.
- Nichols, P., Pokorny, R., Jones, G., Gott, S.P., & Alley, W.E. (in preparation). *Evaluation of an avionics troubleshooting tutoring system*. (Technical Report). Brooks AFB, TX: Armstrong Laboratory, Human Resources Directorate.
- Palincsar, A., & Brown, A. (1984). Reciprocal teaching of comprehension fostering and monitoring activities. *Cognition and Instruction*, 1(2), 117-175.
- Perkins, D.N. (May 1991). Technology meets constructivism: Do they make a marriage? *Educational Technology*, 18-23.
- Polson, M.C., & J. J. Richardson. (1988). *Foundations of intelligent tutoring systems*. Hillsdale NJ: Lawrence Erlbaum.
- Resnick, L.B. (1987) *Learning in school and out*. Presidential address at the Annual Conference of the American Educational Research Association: Washington, DC.
- Scardamalia, M., & Bereiter, C. (1985). Fostering the development of self-regulation in children's knowledge processing. In S.F. Chipman, J.W. Segal, & R. Glaser (Eds.), *Thinking and learning skills: Research and open questions*. Hillsdale, NJ: Erlbaum.
- Schoenfeld, A.H. (1985). *Mathematical problem solving*. New York: Academic Press.
- Simon, H.A. (1981). *The sciences of the artificial*. Cambridge, MA: MIT Press.
- Thorndike, E.L. (1903). *Educational psychology*. New York: Lemke & Buechner.
- Thorndike, E.L. (1906). *Principles of teaching*. New York: A.G. Seiler.
- Thorndike, E.L., & Woodworth, R.S. (1901). The influence of improvement in one mental function upon the efficiency of other functions. *Psychological Review*, 8, 247-261

- Wenger, E. (1987). *Artificial intelligence and tutoring systems*. Los Angeles: Morgan Kaufmann Publishers, Inc.
- Wertheimer, M. (1945). *Productive thinking*. New York: Harper & Row.
- White, B.Y., & Frederiksen, J.R. (1986). *Progressions of qualitative models as foundations for intelligent learning environments*. (BBN Report No. 6277). Cambridge, MA: BBN Laboratories Inc.
- White, B. Y., & Frederiksen, J.R. (1987) *Causal model progressions as a foundation for intelligent learning environments*. (BBN Report No. 6686). Cambridge, MA: BBN Laboratories Inc.